Received: 15 April 2020 | Accepted: 16 May 2020

DOI: 10.1111/gcbb.12723

ORIGINAL RESEARCH

GCB-BIOENERGY BIOPRODUCTS FOR A SUSTAINABLE BIOECONOMY Open Access

WILEY

Genetic improvement of the chemical composition of Scots pine (*Pinus sylvestris* L.) juvenile wood for bioenergy production

Tomáš Funda[1,2,3] | **Irena Fundová**[1,4] | **Anders Fries**[1] | **Harry X. Wu**[1,5,6]

[1]Umeå Plant Science Centre, Department of Forest Genetics and Plant Physiology, Swedish University of Agricultural Sciences, Umeå, Sweden

[2]Department of Genetics and Breeding, Faculty of Agrobiology and Natural Resources, Czech University of Life Sciences, Prague, Czech Republic

[3]Key Laboratory of Forest Genetics and Biotechnology, Co-Innovation Center for the Sustainable Forestry in Southern China, Nanjing Forestry University, Nanjing, China

[4]Skogforsk (Forestry Research Institute of Sweden), Sävar, Sweden

[5]Beijing Advanced Innovation Center for Tree Breeding by Molecular Design, Beijing Forestry University, Beijing, China

[6]National Research Collections Australia, CSIRO, Canberra, ACT, Australia

Correspondence
Tomáš Funda and Harry X. Wu, Umeå Plant Science Centre, Department of Forest Genetics and Plant Physiology, Swedish University of Agricultural Sciences, Umeå, Sweden.
Email: Tomas.Funda@slu.se (T. F.); Harry.Wu@slu.se (H. X. W.)

Funding information
Svenska Forskningsrådet Formas, Grant/Award Number: 2014-00171

Abstract

Chemical composition is one of the key characteristics that determines wood quality and in turn its suitability for different end products and applications. The inclusion of chemical compositional traits in forest tree improvement requires high-throughput techniques capable of rapid, non-destructive and cost-efficient assessment of large-scale breeding experiments. We tested whether Fourier-transform infrared (FTIR) spectroscopy, coupled with partial least squares regression, could serve as an alternative to traditional wet chemistry protocols for the determination of the chemical composition of juvenile wood in Scots pine for tree improvement purposes. FTIR spectra were acquired for 1,245 trees selected in two Scots pine (*Pinus sylvestris* L.) full-sib progeny tests located in northern Sweden. Predictive models were developed using 70 reference samples with known chemical composition (the proportion of lignin, carbohydrates [cellulose, hemicelluloses and their structural monosaccharides glucose, mannose, xylose, galactose, and arabinose] and extractives). Individual-tree narrow-sense heritabilities and additive genetic correlations were estimated for all chemical traits as well as for growth (height and stem diameter) and wood quality traits (density and stiffness). Genetic control of the chemical traits was mostly moderate. Of the major chemical components, highest heritabilities were observed for hemicelluloses (0.43–0.47), intermediate for lignin and extractives (0.30–0.39), and lowest for cellulose (0.20–0.25). Additive genetic correlations among chemical traits were, except for extractives, positive while those between chemical and wood quality traits were negative. In both groups (chemical and wood quality traits), correlations with extractives exhibited opposite signs. Correlations of chemical traits with growth traits were near zero. The best strategy for genetic improvement of Scots pine juvenile wood for bioenergy production is to decrease and stabilize the content of extractives among trees and then focus on increasing the cellulose:lignin ratio.

KEYWORDS

biomass, cellulose, extractives, forest tree breeding, lignin, wood quality

1 | INTRODUCTION

Wood is a natural organic material that has been utilized by humans for millennia in many aspects of their daily lives. Its abundance, versatility and environmental sustainability makes it the material of choice for a broad range of purposes in both raw (fuel, construction timber, furniture, tools, fence posts, utility poles) and processed (paper, textile fabrics, bioethanol) forms.

Wood is composed of four major chemical components: cellulose, hemicelluloses, lignin, and extractives (Poletto, Zattera, & Santana, 2012). The former three, collectively called lignocellulose, are polymeric macromolecules that constitute wood's structural components. Cellulose microfibrils together with hemicellulosic chains that are strong in tension are embedded in a matrix of lignin that resists compression. In conifer tree species, they commonly represent 40%–50%, 25%–35% and 18%–35% of the total dry weight of stem wood, respectively, jointly making up 90%–96% (Pettersen, 1984; Stevanovic, 2016). Wood extractives, forming the rest of wood materials, represent a large and heterogeneous group of low-molecular-weight organic and inorganic compounds (Ekeberg, Flaete, Eikenes, Fongen, & Naess-Andresen, 2006; Fengel & Wegener, 1989), many of which also enjoy commercial utilization in a number of industrial areas (Nisula, 2018). Along with anatomical structure, the chemical composition of wood is the main determinant of wood mechanical properties and, in turn, of wood quality.

The term "wood quality" does not have a single definition, but it can be simplified to properties that reflect the degree of excellence of wood in one way or another, depending on its intended utilization (Barnett & Jeronimidis, 2003). For instance, it will be mainly characterized by stiffness, strength, density and the presence and extent of reaction wood in construction lumber (Fundova, Funda, & Wu, 2018, 2019; Ramage et al., 2017), while fiber length and the chemical composition such as the ratio of carbohydrates to lignin and other wood components will be more important in pulp and paper industries and biofuel production (Wegner, Skog, Ince, & Michler, 2010).

In the context of rapidly changing global climate due to the unprecedented overproduction of greenhouse gas emissions, biofuel from wood is becoming an appealing carbon-neutral energy resource that could potentially replace the consumption of fossil fuels in the future. Biofuels have been predominantly produced from plant species such as corn, sugarcane and soybean, whose tissues contain high proportions of starch and/or oil (Schubert, 2006). However, lignocellulosic biomass produced by forest trees has been receiving considerable interest in this regard too (Binder & Raines, 2009; Pandey, Larroche, Ricke, Dussap, & Gnansounou, 2011) and is forecast to play an important role in satisfying the rising demand for transportation fuels (Sticklen, 2010). To keep pace with the demand brings about the need to generate satisfactory amounts of wood with suitable properties, in particular with suitable chemical composition.

Provided that chemical compositional traits are heritable and exhibit sufficient genetic variation, they can be included in forest tree breeding programs as target traits and improved via recurrent selection. However, the target traits need to be correlated with some proxy (selection) traits that can be rapidly, inexpensively and non-destructively measured on standing trees, preferably at young ages, because standard wet chemistry protocols (SPPBTC, 2003, 2009; TAPPI, 1991, 2002) are laborious and expensive and thus cannot be applied for large-scale evaluations of breeding populations.

Infrared (IR) spectroscopic techniques belong to the most promising alternatives in this context (Gebreselassie et al., 2017; Hein & Chaix, 2014; Lepoittevin et al., 2011; Pot et al., 2002; Schimleck, 2008). In particular, near-infrared (NIR) and Fourier-transform infrared (FTIR) spectroscopies (reviewed by Conrad & Bonello, 2016; Xu, Yu, Tesso, Dowell, & Wang, 2013), have been successfully applied for rapid screening of the chemical composition of wood in a number of forest tree species (Cozzolino, 2014). Here, only a small subset of samples (albeit as representative of the population under study as possible) undergoes the accurate and precise determination of the chemical composition in the wet lab while the remaining samples' composition will be predicted from their IR spectral profiles using multivariate regression modeling (Cozzolino, 2014; SAS, 2008; Zhou, Jiang, Cheng, & Via, 2015). FTIR spectroscopy is highly sensitive and accurate and provides information-rich spectra with a number of sharp peaks (Faix, 1992; Hergert, 1971), many of which can be directly related to the presence of one or more chemical components in question. It thus appears to be an ideal alternative to the traditional wet chemistry protocols, especially when limited variation in the chemical composition is anticipated among samples.

The chemical composition of wood follows a general pattern but varies at both inter- and intra-specific levels as well as among different parts and/or age classes of the same trees (Pettersen, 1984; Räisänen & Athanassiadis, 2013). One significant source of variation is due to different developmental stages of trees, when juvenile and mature woods are produced. In conifer species, the transition between them typically occurs between 10 and 20 years of age (Hayatgheibi et al., 2018; Jozsa & Middleton, 1994), but both the starting point and duration are highly variable and may differ among wood traits even at the same tree (Hodge & Purnell, 1993; Yang & Benson, 1997).

Mature wood is more desirable because it is denser, has longer tracheids and contains less lignin and extractives

(Loo, Tauer, & McNew, 1985; Sykes, Isik, Li, Kadla, & Chang, 2003). However, its proportion at harvest gradually declines because forest tree improvement—at least until recently—primarily focused on increasing stem volume (Wilhelmsson & Andersson, 1993) while traits related to wood quality (and to the transition age from juvenile to mature wood) were not considered. Consequently, improved trees would reach merchantable dimensions sooner and could thus be harvested earlier than their unimproved counterparts (Petty, MacMillan, & Steward, 1990; Zhou & Smith, 1991) but the transition age remained unchanged. For instance, rotation periods of loblolly and radiata pines (*Pinus taeda* L. and *Pinus radiata* D. Don) in the US and Australia, respectively, were shortened by about 50% to 20–30 years compared with natural stands (Gapare, Wu, & Abarquez, 2006; Pearson & Gilmore, 1980), thus leaving trees with too little time for producing mature wood.

Large genetic variation in the chemical compositional traits has been observed in juvenile wood of several pine species (Shupe, Choong, & Yang, 1996; Sykes, Li, Isik, Kadla, & Chang, 2006). Since these traits are directly linked with the usability of wood for different applications, including pulp, paper and biofuel production, it is appealing to incorporate them in forest tree breeding programs. In this study we intended to (a) quantify the extent of additive genetic variation in growth, wood quality and chemical compositional traits in juvenile wood of Scots pine (*Pinus sylvestris* L.), (b) estimate all traits' narrow-sense heritabilities, (c) determine the magnitude and direction of phenotypic and additive genetic correlations between all pairs of the studied traits and (d) evaluate the potential of the chemical compositional traits for genetic improvement via selective breeding to produce wood materials with desired chemical compositions.

2 | MATERIALS AND METHODS

2.1 | Sample population

Samples for this study were selected in two Scots pine (*Pinus sylvestris* L.) full-sib progeny tests "Skorped" (411-2-H72-Skorped-Y; latitude 63.3444°N, longitude 17.6417°E, altitude 330 m a.s.l.) and "Vännäsby" (411-3-V73-Vännäsby-AC; latitude 64.0250°N, longitude 19.8519°E, altitude 200 m a.s.l.). Seeds for both tests were sown in May 1972 at Skogforsk (Forestry Research Institute of Sweden), Sävar. The tests were established in October 1972 and May 1973, respectively, on normal forest soils with intermediate fertility using completely randomized single tree plot design.

The tests were part of a broader progeny test series and consisted of 199 and 197 controlled crosses, respectively, from seed orchard #411 "Domsjöägnet", three controlled crosses from other seed orchards, six seed stand seedlots and three provenances of lodgepole pine (*Pinus contorta* Douglas ex Loudon), all planted as 1 year old seedlings in paper pots. The test sites were divided into 210 and 208 postblocks, respectively, each consisting of 40 trees (4 columns by 10 rows with spacing of 2.2 m in each direction). The total number of planted trees was 8,390 and 8,320 on the two sites. A subset of 1,245 trees, representing 105 full-sib families (629 and 616 trees, respectively) that were planted at both sites, was included in this study. At each site, 85 families were represented by at least five trees while 20 families were represented by at least 10 trees, provided that enough surviving trees existed on the sites to meet this condition.

2.2 | Traits

2.2.1 | Growth traits

Height (*HGT*) and diameter at breast height (*DBH*) at age 28 were obtained from Skogforsk who had measured the plantations in the fall of 2000. Merchantable volume (*VOL*) was calculated as a function of height and diameter following Brandel (1990) as:

$$VOL = \exp\left[-2.7841 + 1.9474 \cdot \ln(DBH) - 0.05947 \cdot \ln(DBH+20) + 1.40958 \cdot \ln(HGT) - 0.4581 \cdot \ln(HGT-1.3)\right]. \quad (1)$$

2.2.2 | Wood quality traits

Density

Wood density (*RES*) was estimated from records obtained in the fall of 2015 using a portable micro-drill Resistograph IML-RESI PD300 (Instrumenta Mechanic Labor), which measures drilling resistance of wood with a resolution of 10 points per mm. Each standing tree was drilled bark to bark c. 1.3 m above ground, with a reasonable distance kept from knots and observable stem damages. The drilling was performed in one direction only (from northeast and west at Skorped and Vännäsby, respectively), as no significant differences had been observed between measurements in two perpendicular directions (Fundova, Funda, & Wu, 2018). When resistograms (drilling resistance values plotted against penetration depths), as immediately displayed on the instrument's screen, exhibited inconsistencies or large deviations from the expected pattern, measurements were repeated. Raw resistograms were adjusted (debarked and linearly detrended to correct for needle friction) following Fundova et al. (2018), and *RES* was subsequently calculated as the mean of all drilling resistance values along a given resistogram.

Dynamic modulus of elasticity

Acoustic velocity (*VEL*) was recorded on standing trees in the fall of 2015 using a portable device Hitman ST300 (Fibregen; Carter, Briggs, Ross, & Wang, 2005). The device records the time of flight of mechanically induced dilatational stress waves between two Monitran MTN/P100 accelerometers that are attached to probes hammered into a tree's stem at c. 45-degree angle. Distance between probes, measured with ultrasonic sensors, was maintained between 0.5 and 1.2 m while the lower probe was situated at c. 0.6 m above ground, depending on the occurrence of knots. Measurements were taken in the same directions as the resistograph's, thus from northeast and west at Skorped and Vännäsby, respectively. Each acoustic velocity value for a given tree, which was supplied into the formula below, was an average of two series of eight successive records that were taken with the aim of accounting for variation among records on the same tree (Paradis, Auty, Carter, & Achim, 2013). When estimates from the two series differed by more than c. 3%, a third series was generated. Wood stiffness, expressed as the dynamic modulus of elasticity (MOE_d; GPa), was calculated following Bucur (2006) as:

$$MOE_d = VEL^2 \cdot \rho, \quad (2)$$

where *VEL* is the acoustic velocity (km/s) and ρ is the wood density (kg/m^3). Unitless resistograph-based density (*RES*) was used as a surrogate for wood density, as it had been reported to be highly correlated with x-ray density, with the phenotypic and additive genetic correlation coefficients reaching 0.72 and 0.96, respectively (Fundova et al., 2018). The original resistograph density values were divided by four so that they were scaled down to approach actual density values expressed in kg/m^3 (Fundova, Hallingbäck, Jansson, & Wu, 2020). Note that the measurements of *VEL* and the subsequent calculation of MOE_d were only conducted at the Vännäsby site.

2.2.3 | Chemical compositional traits

In spring 2016, bark to bark increment cores were extracted from standing trees at breast height using a 5 mm core borer (Haglöf), powered by a battery-operated portable device, and were stored in paper straws at c. +23°C and 30% relative humidity. The extraction followed the drilling trajectory of the resistograph's measurements performed in the previous season, with sufficient distance (5–10 cm) maintained between them to avoid their crossing. Annual rings 2–6 counted from the pith, representing juvenile wood and corresponding to tree age of c. 9–13 years, were subsequently isolated for chemical and FTIR analyses. Wood samples were ground using a Retch MM400 ball mill (Retch GmbH) in two cycles by 40 s each, with a 2 min gap between them to avoid overheating, and either used as such for chemical analyses or manually fine-ground with IR spectroscopy grade KBr (Sigma-Aldrich) to produce a homogenized mix for FTIR analyses. The manual grinding was performed using an agate pestle and mortar at a weight ratio of 1 unit of wood powder to c. 55 units of KBr (Gorzsás & Sundberg, 2014).

The chemical analyses were performed at MoRe Research (Örnsköldsvik) on a subset of 70 trees (34 from Vännäsby and 36 from Skorped), which were selected with the aim of covering as much phenotypic variation as possible in wood density and stiffness as estimated earlier with the resistograph and Hitman, respectively. The content of carbohydrates was determined following the protocol SCAN-CM 71:09 (SPPBTC, 2009). Monosaccharides glucose, xylose, mannose, galactose, and arabinose, denoted hereafter as *GLU*, *XYL*, *MAN*, *GAL*, and *ARA*, respectively, were quantified by Dionex ICS-5000 ion chromatography (Thermo Scientific Inc.). The ratio of cellulose (*CEL*) to hemicelluloses (*HEM*) was derived from the relative content of *GLU* and *MAN* following the formula developed by Sjöström (1993) for softwoods as $CEL = GLU - MAN/3$ and $HEM = 1 - CEL$. Total lignin (*LIG*) was quantified as the sum of Klason and acid-soluble lignin following TAPPI's protocols 222 om-02 (TAPPI, 2002) and UM 250 (TAPPI, 1991), respectively. Analysis of extractives (*EXT*) followed the Soxhlet extraction protocol SCAN-CM 67:03 (SPPBTC, 2003), with a 9:1 ratio of cyclohexane:acetone. The original protocols were slightly modified to account for low amounts of wood material (c. 200 mg per sample).

FTIR measurements of all 1,245 juvenile wood samples included in this study were performed at the Vibrational Spectroscopy Core Facility of Umeå University using a Bruker IFS 66v/S vacuum bench spectrometer (Bruker Optics) in a 16-unit automatic diffuse reflectance carousel (Harrick Scientific Products Inc.). Spectra were collected over the range of 5,200–400 cm^{-1} at 4 cm^{-1} spectral resolution, employing a zero filling factor of 2 and Blackman-Harris three-term apodization function. Each sample was scanned 128 times in order to attain good signal to noise ratios. Raw spectra were exported using OPUS 7.0 (Bruker Optics) and standardized for subsequent prediction model calibration using an open source graphical user interface available at https://www.umu.se/en/research/infrastructure/visp/. The standardization comprised of IR wave spectra trimming to only retain the so-called "fingerprint" wave region of 1,869–771 cm^{-1}, baseline correction via asymmetrical least squares fitting (Eilers, 2004), normalization using either total area (TAN) or area minimum-maximum (AMM1 and AMM2) normalization, and smoothing following Savitzky–Golay filtering (Savitzky & Golay, 1964).

2.3 | Model development

The chemical composition of all 1,245 juvenile wood samples was predicted using partial least squares regression (PLSR) models developed by Funda, Fundova, Gorzsás, Fries, and Wu (2020) in SAS 9.4 (SAS Institute Inc.) based on spectral and chemical compositional data obtained for the selected 70 samples. Standardized FTIR spectra within 771–1,869 cm^{-1} served as predictor variables (in total 570 variables; each representing absorbance intensity at a given wavenumber) while the nine chemical compositional traits *LIG*, *CEL*, *HEM*, *GLU*, *MAN*, *XYL*, *GAL*, *ARA*, and *EXT*, expressed as percentages of the total dried wood content, served as response variables.

Each response variable was modeled separately, and the models were validated using a split-sample cross-validation test, in which groups of every seventh observation were excluded from calibration data sets. The normalized root mean squared error of predictions (*RMSEP*), provided by SAS 9.4, was used as the benchmark statistics during calibration. Vandervoet's randomization-based model comparison test (Vandervoet, 1994) was applied as the primary criterion for model selection.

2.4 | Quantitative genetic analyses

All statistical analyses related to the estimation of quantitative genetic parameters were performed in the statistical package ASReml 4 (VSN International Ltd.). The two progeny tests Vännäsby and Skorped were analyzed as two separate single-site analyses and, with all data combined, also in multisite analysis. All growth, wood quality and chemical compositional traits (in total 14 variables) were fitted into the following linear mixed model.

$$y_{ijklm} = \mu + g_i + g_j + s_{ij} + t_l + p_k\left(t_l\right) + e_{ijklm}, \quad (3)$$

where y_{ijklm} is the mth individual for an offspring of ith and jth parents growing in the kth plot on the lth site, μ is the overall mean of a given response variable, g_i and g_j are the random general combining ability (GCA) effects of ith and jth parents, respectively, s_{ij} is the random specific combining ability (SCA) effect for the cross between parents i and j (Griffing, 1956), t_l is the fixed effect of the lth site, $p_k\left(t_l\right)$ is the random effect of the kth plot nested within the lth site and e_{ijklm} is the random error term, specific to the $ijklm$th individual. In the single-site analysis, the site-specific terms t_l and $p_k\left(t_l\right)$ were replaced with p_k, which is the random effect of the kth plot. Individual-tree narrow-sense heritabilities for each response variable were estimated using variance components obtained from univariate analyses as:

$$h_i^2 = \frac{\sigma_A^2}{\sigma_P^2} = \frac{4\sigma_g^2}{2\sigma_g^2 + \sigma_s^2 + \sigma_e^2}, \quad (4)$$

where σ_g^2, σ_s^2 and σ_e^2 denote GCA, SCA and residual variance, respectively, and the numerator and denominator represent additive genetic and phenotypic variances. Standard errors of the heritability estimates were calculated following Taylor series expansion incorporated in the ASReml software (Gilmour, Gogel, Cullis, Welham, & Thompson, 2015). Phenotypic and genetic correlations were calculated as:

$$r_{xy} = \frac{\sigma_{xy}}{\sqrt{\sigma_x^2 \times \sigma_y^2}}, \quad (5)$$

where σ_x^2 and σ_y^2 are the phenotypic or additive genetic variances for traits x and y, respectively, and σ_{xy} is the phenotypic or additive genetic covariance between the traits estimated by fitting a bivariate mixed model (Gilmour et al., 2015). Expected genetic gain for direct selection (G_{A_x}) was calculated as:

$$G_{A_x} = ih_{ix}\sigma_{A_x}, \quad (6)$$

where i is the selection intensity (i = 2.665) and h_{ix} and σ_{A_x} are the square roots of the narrow-sense heritability and additive genetic standard deviation for trait x, respectively. The correlated response of the target trait y (CR_y) due to selection for trait x was calculated as:

$$CR_y = ih_{ix}\sigma_{A_y}r_{Axy}, \quad (7)$$

where σ_{A_y} is the additive genetic standard deviation for the target trait y and r_{Axy} is the additive genetic correlation between the selection trait x and the target trait y.

3 | RESULTS

3.1 | Predictive PLSR models

Only one model that performed best in terms of *RMSEP* was retained for a given response variable (Table 1). The overall predictive power of the models was good, with *RMSEP* values ranging from 0.302 for *EXT* to 0.812 for *ARA* (average *RMSEP* for all models was 0.613). With the exception of *ARA* (R^2 = .679), the explained response variation exceeded 75%. The number of significant factors retained following Vandervoet's test ranged from 1 to 9 (Table 1). FTIR spectra standardized following TAN and AMM gave rise to most accurate models in five and four response variables, respectively, although the differences in the attained *RMSEP*s between the best and second best models were minor, on average less than 2%. Following outlier analysis, up to four individuals were removed from the calibration data set. A summary of the

TABLE 1 Performance of partial least squares regression models for predicting the chemical composition of juvenile wood from standardized FTIR spectra (570 predictor variables representing absorbance intensities at 1,869–771 wavenumbers)

Chemical component	Normalization type	Removed outliers	Min *RMSEP*	Factors[a]	*RMSEP*	R^2_x	R^2_y
LIG	AMM2	2	0.445	1 (6)	0.476	.485	.811
CEL	TAN	0	0.715	7 (7)	0.715	.840	.806
HEM	AMM1	4	0.688	5 (7)	0.719	.828	.778
GLU	TAN	2	0.661	8 (8)	0.661	.860	.845
MAN	TAN	1	0.603	6 (10)	0.617	.810	.815
XYL	TAN	1	0.713	8 (10)	0.731	.855	.809
GAL	AMM1	1	0.472	9 (11)	0.483	.929	.934
ARA	AMM1	1	0.754	5 (7)	0.812	.857	.679
EXT[b]	TAN	0	0.277	4 (9)	0.302	.725	.953

Abbreviations: AMM1 & AMM2, area minimum-maximum normalization according to spectral regions 925–1,145 cm^{-1} and 1,487–1,553 cm^{-1}, respectively; R^2_x and R^2_y, predictor and response variation explained; *RMSEP*, root mean squared error of predictions; TAN, total area normalization.

[a]Number of significant factors retained following Vandervoet's randomization-based model comparison test (number of factors corresponding to minimum *RMSEP* in parentheses).

[b]Calibrated using 69 samples.

TABLE 2 Descriptive statistics of the chemical composition (% of total dried wood content) of juvenile wood determined using wet chemistry analyses (based on 70 samples, left) and predicted from standardized FTIR spectra (based on 1,245 samples, right)

Content (%)	Mean	Min	Max	*SD*	*CV*	Mean	Min	Max	*SD*	*CV*
Component	**Laboratory determination**					**FTIR-based prediction**				
LIG	27.1	20.6	32.3	2.26	8.4	26.2	9.9	34.2	2.9	11.0
CEL	32.6	27.1	37.1	2.31	7.1	32.0	22.8	41.5	2.5	7.9
HEM	22.9	18.9	26.9	1.61	7.0	22.2	13.8	27.8	1.9	8.6
GLU	35.3	29.4	40.4	2.51	7.1	34.7	23.6	45.8	2.9	8.2
MAN	7.9	6.1	10.0	0.81	10.2	7.9	5.0	11.7	0.8	10.5
XYL	5.9	4.7	7.4	0.53	8.9	5.8	3.8	7.1	0.5	8.4
GAL	4.7	2.2	9.2	1.39	29.8	4.6	0.2	10.1	1.5	32.8
ARA	1.8	1.2	2.4	0.22	12.3	1.7	0.5	2.3	0.2	12.6
EXT	7.4[a]	2.0[a]	19.8[a]	3.99[a]	53.6[a]	9.2	0.8	29.6	5.2	56.3
Unassigned	10.1	2.8	14.8	2.4	23.8	10.8	6.4	23.9	1.8	17.1

Abbreviations: *CV*, coefficient of variation; *SD*, standard deviation.

[a]Based on 69 samples.

performance of all models used in this study is provided in Table 1.

3.2 | Analysis of the chemical composition

The chemical composition (Table 2) was highly variable among the 70 sampled trees included in wet chemistry analyses. The major chemical components *LIG*, *CEL*, *HEM*, and *EXT* were determined to constitute on average 27.1%, 32.6%, 22.9%, and 7.4% of total dry weight, respectively. The highest variability was observed in *EXT*, which ranged from 2.0% to 19.8%, with the coefficient of variation (*CV*) exceeding 50%. The sum of the total assigned content approached 90%; the unassigned portion that includes mainly pectin, proteins and inorganic compounds ranged from 2.8% to 14.8%.

Considerable variation in the chemical composition was also observed among the 1,245 predictions. While the means and standard deviations were comparable with those obtained from the wet chemistry analyses, the value ranges were greater among the predictions in all nine chemical traits, in particular in *EXT*, *LIG*, *GLU*, and *CEL* that had the range 0.8%–29.6%, 9.9%–34.2%, 23.6%–45.8%, and 22.8%–41.5%, respectively. *EXT* also exhibited the greatest phenotypic variation of all traits (*CV* = 56.3%). Descriptive statistics of the chemical composition determined

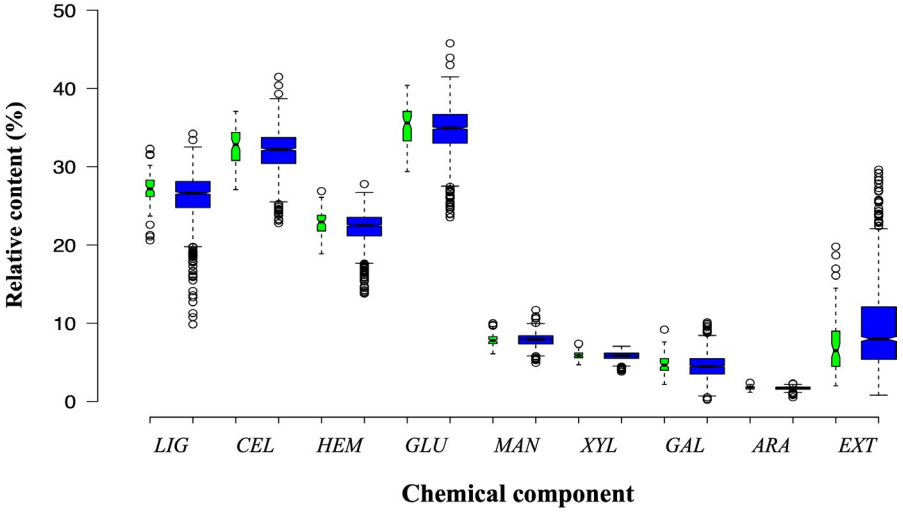

FIGURE 1 Descriptive statistics of the chemical composition of juvenile wood determined using wet chemistry analyses (green boxes; 70 samples) and predicted from standardized Fourier transform infrared spectra (blue boxes; 1,245 samples). Boxes were constructed following Tukey's method. *ARA*, arabinose; *CEL*, cellulose; *EXT*, extractives; *GAL*, galactose; *GLU*, glucose; *HEM*, hemicelluloses; *LIG*, lignin; *MAN*, mannose; *XYL*, xylose

TABLE 3 Descriptive statistics of growth and wood quality traits of Scots pine progeny tests Vännäsby and Skorped

Trait	Unit	*n*	Mean	Min	Max	*SD*	*CV*
DBH[a]	cm	1,214	16.0	8.3	26.6	3.1	19.2
HGT[a]	m	1,245	10.1	5.6	12.9	1.0	9.7
VOL[a]	dm^3	1,214	111.7	21.7	305.2	45.9	41.1
RES	—	1,243	485.6	275.5	797.2	55.2	11.4
VEL[b]	km/s	609	4.0	2.7	4.7	0.3	7.6
MOEd[b,c]	GPa	609	8.1	3.5	12.4	1.7	20.7

Abbreviations: *CV*, coefficient of variation; *n*, sample size; *SD*, standard deviation.

[a]Measured at age 28 (fall 2000).

[b]Measured at Vännäsby only.

[c]Calculated from scaled density values provided by resistograph IML Resi PD-300.

from wet chemistry analyses and predicted from FTIR spectra is visualized in Figure 1 and summarized in Table 2.

3.3 | Growth and wood quality traits

Descriptive statistics of growth and wood quality traits for all 1,245 trees are shown in Table 3. Substantial variation was observed for *DBH* (range 8.3–26.6 cm; *CV* = 19.2%) and especially for *VOL* (21.7–305.2 dm^3; *CV* = 41.1%). Wood quality traits exhibited some variation too, with *VEL* the least (2.7–4.7 km/s; *CV* = 7.6%) and MOE_d the greatest (3.5–12.4 GPa; *CV* = 20.7%). Note that MOE_d estimates shown in Table 3 were calculated from density values provided by the resistograph and scaled down to 25%; therefore, they might lie slightly above or below true values, depending on the actual scaling coefficient.

3.4 | Genetic variation and narrow-sense heritabilities

Coefficients of additive genetic variation (CV_A) for growth, wood quality and chemical compositional traits along with their narrow-sense heritabilities (h_i^2), estimated separately for each site as well as jointly for both sites, are presented in Table 4.

CV_A exhibited a consistent pattern across traits at the two sites, with values ranging from 2.7% for *MAN* at Vännäsby to 36.1% for *EXT* at Skorped. *VOL*, MOE_d and *EXT* exhibited the highest CV_A within the respective trait groups while *HGT*, *VEL* and *CEL* & *GLU* exhibited the lowest.

Assessed based on the magnitude of standard errors (Porth et al., 2013), all heritabilities were significant. Heritabilities for growth traits ranged from 0.16 for *DBH* to 0.38 for *HGT* for the single site analyses (both at Skorped), and from 0.10 also for *DBH* to 0.33 for *HGT* in the multisite analysis. Wood quality traits exhibited the highest heritabilities, exceeding 0.5 for the three traits *RES*, *VEL* and MOE_d at Vännäsby, with *RES* reaching appreciable levels also at Skorped and on both sites combined, 0.30 and 0.42, respectively. The genetic control of chemical compositional traits was mostly moderate and, except for *MAN* and *GAL*, also consistent between sites. These two traits differed in their heritabilities between sites by 0.19 and 0.17, respectively, with higher values being attained consistently at Skorped; the remaining

TABLE 4 Coefficients of additive genetic variation (CV_A) and individual narrow-sense heritabilities (h_i^2) with standard errors in parentheses

	CV_A (%)	h_i^2	CV_A (%)	h_i^2	CV_A (%)	h_i^2
	Vännäsby		**Skorped**		**Combined**	
DBH	7.5	0.20 (0.08)	6.9	0.16 (0.07)	5.5	0.10 (0.04)
HGT	4.9	0.31 (0.10)	5.7	0.38 (0.10)	5.1	0.33 (0.07)
VOL	16.9	0.21 (0.08)	18.3	0.24 (0.08)	14.8	0.16 (0.05)
RES	7.7	0.55 (0.12)	5.5	0.30 (0.08)	6.6	0.42 (0.09)
VEL	5.5	0.53 (0.12)	N/A	N/A	N/A	N/A
MOE_d	16.0	0.58 (0.13)	N/A	N/A	N/A	N/A
LIG	5.2	0.34 (0.10)	5.6	0.30 (0.09)	5.5	0.32 (0.07)
CEL	3.2	0.20 (0.09)	4.0	0.25 (0.09)	3.8	0.25 (0.06)
HEM	5.1	0.47 (0.12)	5.1	0.43 (0.10)	5.1	0.46 (0.09)
GLU	2.9	0.16 (0.08)	4.1	0.23 (0.09)	3.8	0.23 (0.06)
MAN	2.7	0.08 (0.07)	5.2	0.27 (0.09)	4.6	0.24 (0.06)
XYL	3.9	0.22 (0.09)	3.7	0.21 (0.08)	3.9	0.24 (0.06)
GAL	11.2	0.15 (0.08)	18.5	0.32 (0.09)	14.8	0.24 (0.07)
ARA	6.3	0.30 (0.09)	7.0	0.35 (0.10)	6.6	0.32 (0.08)
EXT	33.7	0.35 (0.10)	36.1	0.39 (0.10)	34.5	0.37 (0.08)

traits reached similar heritability values, differing on average only by 0.04. Of the four major chemical components, highest heritabilities were observed for *HEM* (0.47, 0.43 and 0.46 for Vännäsby, Skorped and the two sites combined, respectively), intermediate for *LIG* and *EXT* (ranging between 0.30 and 0.39) and lowest for *CEL*, only ranging between 0.20 and 0.25. Standard errors were overall reasonably low, reaching on average 36%, 30% and 24% of the heritability estimates.

3.5 | Genetic and phenotypic correlations

Additive genetic (r_A) and phenotypic (r_P) correlations between all pairs of traits are presented in Tables S1–S3. Mostly positive genetic correlations were observed between traits within all three trait groups, i.e., growth, wood quality and chemical compositional traits, and this pattern was consistent across sites as well as in the multisite analysis. The only exceptions to this pattern were strongly negative correlations of *EXT* with all other chemical compositional traits, ranging from −0.60 to −1.07, and close-to-zero correlations of *MAN* and *XYL* with *GAL* at Skorped and *MAN* with *GAL* in the multisite analysis. Within the group of wood quality traits, a strongly positive genetic correlation of 0.81 was observed between *RES* and MOE_d at Vännäsby; however, this estimate might be inflated, as MOE_d was calculated from resistograph density; that between independently measured *RES* and *VEL* was moderate (+0.53).

Genetic correlations between traits across groups showed variable patterns. Those between growth and wood quality traits were either negative (*DBH* and *VOL* with *RES*, *VEL* and MOE_d at Vännäsby) or near zero (all other instances). Those between growth and chemical compositional traits were mostly negligible, with only a few exceptions at Vännäsby (*DBH*, *HGT* and *VOL* with *LIG*, and *HGT* also with *GLU* and *CEL*), Skorped (*HGT* with *HEM*, *MAN* and *XYL*, and all growth traits with *GAL* and *ARA*) and in the multisite analysis (all growth traits with *GAL*), which were significant in either direction. An interesting pattern was revealed for relationships between chemical compositional traits and wood quality traits. *RES*, which was measured at both sites, was either negatively correlated (six, four and six trait pairs for the three models, respectively) or uncorrelated with eight of the nine chemical compositional traits (this group included *LIG* and all of the carbohydrate traits), while it was in all instances positively correlated with *EXT* ($r_A = 0.45$, 0.35 and 0.40, respectively). Furthermore, the two wood-stiffness related traits *VEL* and MOE_d that were measured at Vännäsby, followed a similar pattern to that of *RES*, reaching negative correlations with five (*VEL*) and six (MOE_d) chemical compositional traits while their correlations with *EXT* were moderately positive (0.42 and 0.51, respectively).

For an easier interpretation, all additive genetic (a) and phenotypic (b) correlation coefficients are visualized in heat maps (Figure 2) following a cluster analysis. All heat maps, referring to Vännäsby (left), Skorped (middle) and the multisite analysis (right), exhibited a similar and strong correlation pattern as described earlier, showing (1) three clear patches of positive correlations comprising of all wood quality traits & *EXT*, all growth traits, and all chemical compositional trait except for *EXT*; (2) no or weak correlations, both positive and negative, comprising of growth traits with wood

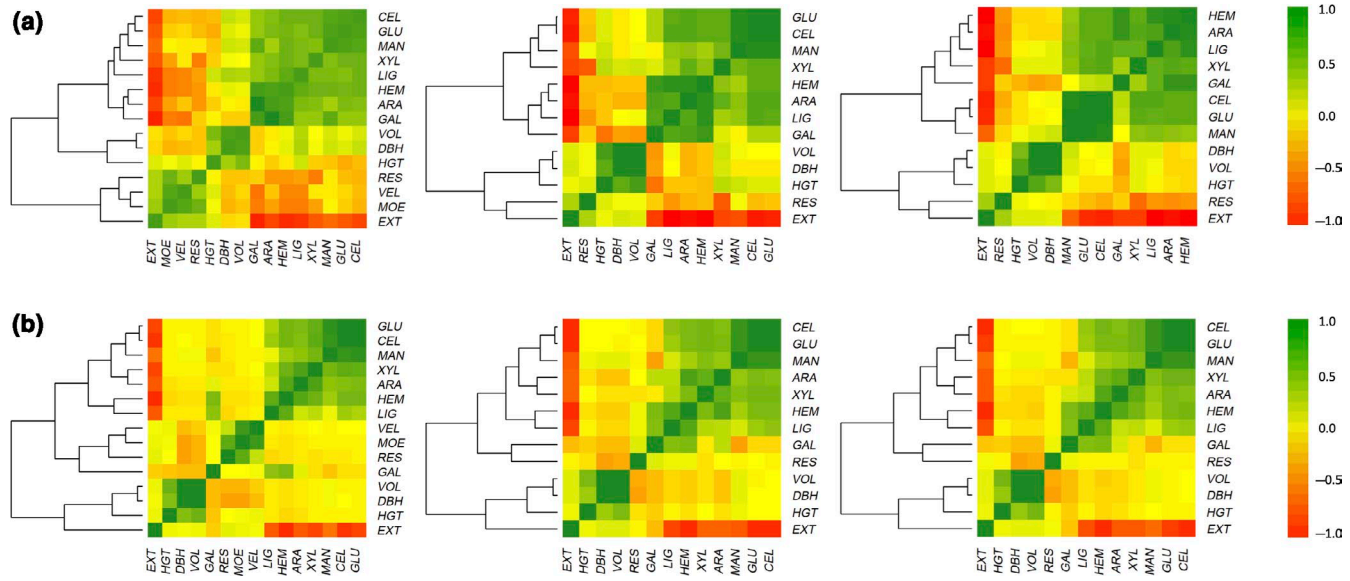

FIGURE 2 Additive genetic (a) and phenotypic (b) correlations of growth, wood quality and chemical compositional traits at research sites Vännäsby (left), Skorped (middle) and for both sites combined via multisite analysis (right). Colors and their shades represent the direction and magnitude of the correlation coefficients, respectively. *ARA*, arabinose; *CEL*, cellulose; *DBH*, stem diameter at breast height; *EXT*, extractives; *GAL*, galactose; *GLU*, glucose; *HEM*, hemicelluloses; *HGT*, tree height; *MAN*, mannose; *MOE*, dynamic modulus of elasticity; *LIG*, lignin; *RES*, resistograph-based wood density; *VEL*, acoustic velocity; *VOL*, strem volume; *XYL*, xylose

TABLE 5 Direct (bold) and correlated response to selection (% of the mean) reflecting 1% selection intensity ($i = 2.665$) based on multisite analysis (progeny tests Vännäsby and Skorped)

	Target trait							
Selection trait	***DBH***	***VOL***	***RES***	***MOE_d***[a]	***CEL***	***LIG***	***HEM***	***EXT***
DBH	**4.68**	12.29	−0.23	−6.56	0.13	0.42	−0.39	3.54
VOL	5.67	**15.83**	−0.07	−5.62	−0.01	0.29	−0.71	5.15
RES	−0.38	−0.26	**11.49**	25.84	−1.96	−3.61	−3.56	23.93
VEL[a]	−4.33	−6.47	7.88	28.52	−1.81	−5.89	−5.67	27.48
MOE_d[a]	−5.28	−9.81	12.76	**32.63**	−2.31	−6.45	−6.53	35.08
CEL	0.59	−0.05	−2.66	−6.75	**5.05**	5.42	4.88	−39.73
EXT	1.06	3.34	4.27	12.94	−5.22	−8.55	−8.02	**55.57**

[a]Based on single site analysis—Vännasby.

quality traits and *EXT* and growth traits with chemical compositional traits except for *EXT*; and (3) one distinct stripe of strong negative correlations, representing the negative relationships between *EXT* and all of the other eight chemical components.

3.6 | Genetic gain and correlated response to selection

Expected genetic gain and correlated genetic response to indirect selection at 1% selection intensity are presented in Table 5. Genetic improvement following direct selection was estimated to range from 5% for *DBH* and *CEL* to 56% for *EXT*, and considerable gains were predicted for *VOL* (16%) and wood quality traits, reaching over 11% for *RES* and 33% for MOE_d.

Selection for either of the wood quality traits (*RES* and MOE_d) would result in a positive and favorable response of the other trait: *RES* would improve stiffness by almost 26% while MOE_d would improve density by 13%. It would at the same time decrease *LIG* by 3.6%–6.5%, but also slightly decrease *CEL* (by c. 2%) and substantially increase *EXT* (by 24%–35%). Selection for *VEL* alone would result in an improvement of both target wood quality traits, with 8% gain attained for density and 29% for stiffness. Selection for growth traits would have a negligible effect on chemical compositional traits and small but unfavorable on MOE_d.

Selection for higher *CEL* would not affect growth traits but it would slightly decrease wood quality traits (*RES* by 3%

and MOE_d by 7%), increase *LIG* and *HEM* (by 5%) and substantially reduce *EXT* (by 40%). The reduction in *EXT* could be even stronger (56%) if selection was targeted against *EXT*, while the increase in *CEL* would be nearly the same. This strategy would, however, result in a higher increase in *LIG* compared to the strategy of targeting for higher *CEL* (c. 3% difference). Furthermore, selection against *EXT* would incur the reduction in wood density (by 4%) and stiffness (by 13%).

4 | DISCUSSION

In forest tree improvement programs, the success of the utilization of IR spectroscopic techniques depends on (a) IR spectra-based predictability of those industrially important mechanical and/or chemical properties, which are to be improved via selective breeding, from IR spectra; (b) extent of genetic variation present in these traits and their genetic control, i.e., their narrow-sense heritabilities; and (c) knowledge of the genetic relationships among these traits.

4.1 | Model reliability

The overall performance of our predictive models was good, with the attained *RMSEP* and R^2_y values conforming to our expectations based on earlier literature.

For cellulose (trait *CEL*): The predictive model for *CEL*, the most important component of lignocellulosic biomass in relation to bioethanol and pulp and paper production, reached an *RMSEP* of 0.72 and R^2_y of .81. By comparison, Acquah, Via, Fasina, and Eckhardt (2016) obtained an R^2 of .72 in their study of loblolly pine biomass acquired from harvesting operations in southern USA, using first-derivative treated FTIR spectra of the fingerprint region, although with a relatively small ratio of performance to deviation (RPD; 1.61). Similar power for predicting this component was reported by Bjarnestad and Dahlman (2002) in a study focusing on the characterization of hardwood and softwood pulps originating from different manufacturing processes. Toivanen and Alen (2006) did not provide a direct estimate for *CEL* in Scots pine, but their *RMSEP* for *GLU* reached 1.7, the highest of all studied traits. Nuopponen, Birch, Sykes, Lee, and Stewart (2006) modeled the chemical composition in Sitka spruce (*Picea sitchensis* (Bong.) Carrière), Scots pine and 24 different hardwood species and obtained *RMSEP*s for *CEL* of 3.3 and 2.8 when most of the mid-infrared (MIR) spectral region and only five principal wavenumbers were included as predictor variables, respectively.

For hemicelluloses (trait *HEM*): Model performance for predicting *HEM* was nearly the same as for *CEL* in this study, with *RMSEP* and R^2_y reaching .72 and .79, respectively; however, this component seems to be more difficult to model, at least when only the fingerprint region (c. 1,800–700 cm^{-1}) is retained for model building. For instance, Acquah et al. (2016) observed that R^2 increased by 10% when full MIR spectra were included, as opposed to the scenario when all non-fingerprint regions had been cut off. This suggests that non-fingerprint regions might encompass some relevant information pertaining to the presence and content of *HEM*, in particular as the differences in R^2 were much smaller for the other three major components. This hypothesis could be supported by Zhou, Jiang, Via, Fasina, and Han (2015), who obtained superior models for *HEM* compared to those for *CEL* when spectral data acquired over the whole MIR range of 4,000–650 cm^{-1} were included, but this hypothesis could not be verified as the fingerprint region alone was not analyzed separately.

For total lignin (trait *LIG*): In this study, *LIG* could be predicted from spectral data with a higher accuracy than either of the two major carbohydrate components. The *RMSEP* for *LIG* was one-third fold lower than that for *CEL* (0.48 vs. 0.72), and this observation was in congruence with several earlier studies (e.g., Nuopponen et al., 2006; Zhou, Jiang, Via, et al., 2015. Accurate models for predicting *LIG* were constructed by He and Hu (2013) in a study involving 147 woody species from China and also by Jiang et al. (2014) in an NIR-based study of pine lumber.

For extractives (trait *EXT*): The model for *EXT* reached the highest predictive power (*RMSEP* = 0.30) of all chemical compositional traits included in this study. This high power was likely driven by a band position near 1,693 cm^{-1}, whose absorbance intensity exhibited the strongest association with this composite trait, with correlation coefficients ranging from 0.91 to 0.97 depending on the normalization method applied (Funda et al., 2020). Furthermore, only less than 5% of the response variation remained unexplained by the model. A high predictive power for *EXT* was also reported by Meder, Gallagher, Mackie, Bohler, and Meglen (1999), Zhou, Jiang, Via, et al. (2015) and Nuopponen et al. (2006) (in the latter study, *EXT* were referred to as wood resin) and, when only the major wood components were taken into account, *EXT* performed superior in Toivanen and Alen (2006), whose model also explained nearly 97% of the response variation.

Based on the above studies, there seems to be a general pattern in the predictability of the four major components from FT-MIR spectra, with lower accuracy being attained for carbohydrates while highest for lignin and, specifically, extractives. As one possible explanation, Acquah et al. (2016) attributed these results to the similar molecular makeup of polysaccharides versus the specific chemical structures of lignin and extractives, giving rise to highly distinctive patterns of IR absorption bands. Moreover, this pattern seems to hold true for NIR spectra as well: for instance, Acquah, Via, Fasina, and Eckhardt (2015) obtained for loblolly pine logging residues the same ranking in the predictability of major

wood components (*EXT* > *LIG* > *CEL* > *HEM*), and Zhou, Jiang, Via, et al. (2015) found similar predictability rankings (*LIG* > *EXT* > *HEM* > *CEL*) for four different hardwood species. Both studies used FT-NIR spectroscopy.

4.2 | Genetic control of growth, wood quality and chemical compositional traits

The genetic control of economically important growth (height, stem diameter, volume) and timber quality (density, stiffness, microfibril angle) traits has been well documented in a number of forest tree species (Chen et al., 2014, 2015; El-Kassaby, Mansfield, Isik, & Stoehr, 2011; Hayatgheibi, Fries, Kroon, & Wu, 2017; Hong, Fries, & Wu, 2014; Isik, Li, & Frampton, 2003; Ivkovich, Namkoong, & Koshy, 2002; Pot et al., 2002; Wu et al., 2008). However, less focus has been dedicated to the study of chemical compositional traits, although these participate in forming the overall quality of wood and in turn in determining its suitability for different end products and industrial applications.

Forest trees generally encompass substantial genetic variation in growth and wood quality traits. Their genetic control is usually rather weak for the former and moderate for the latter. Narrow-sense heritabilities (h_i^2), similar to those obtained in this study, have been reported for Scots pine's growth traits (Fries, 2012; Fundova et al., 2018; Hong et al., 2014) and slightly higher for Scots pine by Fundova et al. (2020) and Haapanen, Velling, and Annala (1997) specifically for *DBH*. In our study, heritability of wood density (*RES*) obtained at Skorped (0.30) was in congruence with that reported by Fries (2012) and Haapanen et al. (1997), whereas estimates closer to the value obtained in the multisite analysis (0.42) were reported by Hong et al. (2014) and Fundova et al. (2018). Heritabilities of *VEL* measured on standing trees as well as of MOE_d were higher compared to those obtained in other studies on Scots pine (Fundova, Funda, & Wu, 2019; Hong et al., 2014).

Genetic control of the chemical compositional traits does not seem to exhibit a consistent pattern across and within species, as their narrow-sense heritabilities have been reported to range from weak to high, with the majority of them being moderate ($0.2 < h_i^2 < 0.5$). Most h_i^2 obtained in this study were moderate, both at the two sites individually and in the multisite analysis, and their standard errors were within a reasonable range of 0.06–0.12, reaching on average 29.6% of the heritability estimates. Only three monomeric sugars, *GLU*, *MAN* and *GAL*, exhibited low h_i^2 at Vännäsby. Somewhat lower estimates were reported by Sykes et al. (2006) in a study of juvenile and transition wood in loblolly pine, using the third and eighth annual rings from the pith at breast height, respectively. Their chemical compositional traits *CEL* and *LIG* exhibited weak individual-tree heritabilities (0.15 and 0.12) and were associated with relatively high standard errors (0.14 and 0.17, respectively). The high errors were attributed to a small number of parents included in the experiment and possible random genetic drift due to sampling bias. A comprehensive study was provided by Pot et al. (2002), who applied FTIR spectroscopy for estimating genetic parameters of chemical compositional traits in maritime pine (*Pinus pinaster* Ait.). They observed moderate to high heritabilities for holocellulose (i.e., total carbohydrates comprising of *CEL* and *HEM*) and *LIG* ($h^2 = 0.47$ and 0.36, respectively); however, after decomposing holocellulose into the two components, only *CEL* exhibited moderate genetic control ($h^2 = 0.34$) while no significant genetic control was detected for *HEM*, which was under the strongest genetic control in this study. Furthermore, Pot et al. (2002) found no genetic control for *EXT*. In another study on maritime pine, in which the chemical composition was predicted from NIR spectra, Lepoittevin et al. (2011) obtained substantially lower narrow-sense heritabilities for *CEL* (0.08) and *LIG* (0.05 and 0.25 at two different sites), but the moderate genetic control of *EXT* (0.35) was in congruence with our estimate as well as with estimates obtained in several other studies (Cown, Young, & Burdon, 1992; Lepoittevin et al., 2011; Zhou, Li, Huang, Chen, & Lin, 2000). Moderate to high genetic control of *EXT* was also reported in Scots pine by Partanen, Harju, Venalainen, and Karkkainen (2011) and Fries, Ericsson, and Gref (2000). Outside the world of conifers, Porth et al. (2013) observed higher h_i^2 for most chemical compositional traits determined through traditional wet lab approaches in a black cottonwood (*Populus trichocarpa* Torr. & Gray) population from western Canada and US. Their estimates were 0.42, 0.35 and 0.65 for *CEL*, *HEM* and *LIG*, respectively, while those for the five monomeric sugars were on average more than twofold higher than those obtained in this study. Standard errors associated with their estimates were however also substantially higher, on average 0.16 versus 0.07 over eight chemical compositional traits included in both studies.

Chemical compositional traits exhibited low genetic variation except for *GAL* and *EXT*, for which the coefficients of additive genetic variation (CV_A) were three and seven times higher compared to other traits within their trait group, respectively. Slightly lower CV_A were observed by Ukrainetz, Kang, Aitken, Stoehr, and Mansfield (2008) and Pot et al. (2002) for some of these traits.

4.3 | Genetic and phenotypic correlations

The knowledge of the magnitude of additive genetic correlations among traits is essential for understanding the correlated response of different traits to selection (Cheverud, 1988), that is, for quantifying how selection for one trait (or a group of traits) affects the performance of another trait. This question is particularly relevant in forest tree improvement, as some economically important traits that influence the overall

economic value of trees or their products such as growth and wood quality traits are often negatively correlated (El-Kassaby et al., 2011; Fries, 2012; Hong et al., 2014; Li & Wu, 2005; Pot et al., 2002; Zhang & Morgenstern, 1995), and thus, their simultaneous genetic improvement might be challenging. In this study, most of the genetic correlations within groups of growth and wood quality traits were positive and significant, although both their direction and magnitude were driven by the mutual dependencies in their estimations, as *VOL* was calculated from *HGT* and *DBH* while MOE_d was calculated from *VEL* and *RES*. Moderate positive genetic correlation between *VEL* and *RES* at Vännäsby was in line with other studies (Fundova et al., 2019, 2020). Correlations between growth and wood quality traits were weakly negative or nonsignificant.

Additive genetic correlations between any two traits among the nine chemical compositional traits were nearly in all cases strong and positive. This is in accordance with expectations for the relationships of *CEL* and *HEM* with their respective monosaccharide constituents (although Porth et al., 2013 reported no significant genetic relationship and significant but weak phenotypic relationship between *CEL* and *GLU* in black cottonwood), but unlike many other studies (Lepoittevin et al., 2011; Pot et al., 2002; Sykes et al., 2003, 2006) as well as Porth et al. (2013), we observed a positive and strong correlation between *CEL* and *LIG* that held for both of the progeny tests included in this study, with genetic correlations ranging from 0.74 to 0.80, while those reported in the aforementioned studies were −0.98, −1.03, −0.99, −0.73 and −0.45, respectively (the first being determined in a clonal trial). One possible explanation for this discrepancy could be the great variation in the content of *EXT* among trees included in this study (range 0.8%–29.6% as predicted from FTIR spectra or 2.0%–19.8% as determined in the wet lab; $CV_P = 56.3\%$ and 53.6%, respectively). Since the chemical compositional traits are expressed in % (or in $g \times kg^{-1}$ of total dry weight), and thus sum up to no more than 100%, one trait (or a group of traits) can only exist at the expense of another trait. In this study, the high variation in *EXT* strongly influenced the remaining components *CEL*, *HEM* and *LIG*, as their value ranges and *CV*s were substantially smaller, and perhaps masked true relationships among them. For instance, Lepoittevin et al. (2011) observed a much lower variation in *EXT* on a site with half-sib progenies compared with our results (mean value = 6.7%, $CV_P = 19.0\%$, $CV_A = 11.3\%$).

The first breeding step for increasing *CEL* in juvenile wood within the studied material therefore seems to be to decrease *EXT* and then focus on improving the *CEL*:*LIG* ratio. Considerably less extractives were observed among the same trees in mature wood (Funda et al., 2020), and thus, mature wood's correlation matrix might exhibit a different pattern, reflecting more the results reported in studies which did not include extractives in their analyses. Sykes et al. (2006) and Porth et al. (2013) obtained negative correlations between *CEL* and *LIG* in their studies of loblolly pine and black cottonwood, respectively, but these variables could not be linked to *EXT* because *EXT* were not determined as a separate trait. Genetic correlations among individual monomeric sugars were all positive and, judging based on the magnitude of their standard errors, in most cases also significant. On the other hand, no clear correlation pattern was observed in black cottonwood (−0.57 for *GAL–ARA* to 0.67 for *MAN–XYL*; Porth et al., 2013) and Douglas-fir (*Pseudotsuga menziesii* (Mirbel) Franco; −1.00 for *GLU–ARA* to 0.51 for *GLU–MAN*; Ukrainetz et al., 2008).

In our study, additive genetic correlations between growth and chemical compositional traits were mostly negligible or very weak. The only somewhat meaningful correlations were observed between growth traits and *LIG* at the Vännäsby site, which were weakly to moderately positive. Similar results based on phenotypic correlations were reported for Douglas-fir by Ukrainetz et al. (2008), who observed that larger trees (with greater height and diameter) contained more lignin and less carbohydrates. Although Ukrainetz et al. (2008) did not study *CEL* individually, the strongly negative genetic correlations observed between growth traits and *GLU* (−0.65 for *HGT* to −0.74 for *VOL*) might suggest a similar relationship between growth traits and *CEL* too, as most of the glucose in Scots pine is present in the form of cellulose while only a smaller portion exists in the hemicelluloses. A similar observation was reported by Pot et al. (2002) in maritime pine, whose genetic correlation between *HGT* and *LIG* was +0.40 while that between *HGT* and *CEL* was −0.37. Costa e Silva (1998) also reported positive genetic correlations between *LIG* and growth traits in a clonal test with Sitka spruce (+0.47 with *DBH* and +0.34 with *HGT*). In this study, growth traits and *CEL* were practically uncorrelated like in Sykes et al. (2006) or Lepoittevin et al. (2011) in loblolly and maritime pines, respectively; however, the above results indicate that there might exist an unfavorable genetic correlation pattern in growth versus chemical compositional traits from the perspective of tree breeding that aims to increase volume production but, at the same time, improve *CEL*:*LIG* ratio in wood.

Genetic correlations between chemical compositional traits and wood quality traits were studied e.g., by Ukrainetz et al. (2008), Pot et al. (2002), Sykes et al. (2003), Porth et al. (2013) and Lepoittevin et al. (2011). Positive genetic correlations between *CEL* and x-ray-based wood density were found by Pot et al. (2002) in maritime pine and Porth et al. (2013) in black cottonwood (+0.62 and +0.74, respectively) as well as by Sykes et al. (2003) in loblolly pine (+0.56), who used volumetric density. In contrast, Ukrainetz et al. (2008) found no significant relationship between *GLU* and x-ray density. Correlations between *LIG* and wood density had mostly the negative sign—moderate estimates of −0.54 and −0.42 were reported by Pot et al. (2002) and Porth et al. (2013), respectively. A weakly negative relationship was also observed between *LIG* and

Pilodyn-based density by Costa e Silva et al. (1998) in a clonal test with Sitka spruce. In this study, weak to moderate negative genetic correlations were obtained for resistograph density (and stiffness-related traits at Vännäsby) with *LIG* as well as with all carbohydrate traits (both poly- and monosaccharides); *EXT* represented the only exception, as they were positively correlated with *RES* on both sites as well as with *VEL* and MOE_d at Vännäsby. This might suggest that the presence of extractives, or at least some of their components, influences either wood density per se or the drill needle penetration during wood density measurements with the resistograph.

4.4 | Genetic gain and correlated response to selection

Genetic gains predicted for the studied traits were small for *DBH* and *CEL*, moderate for *VOL* and *RES* and high for MOE_d and *EXT*. The correlated genetic response to indirect selection showed that selection for higher *CEL* would result in increased *LIG*, as these two components follow the same trajectory due to the positive and strong additive genetic correlation between them. Similarly, it would result in increased *HEM*. On the other hand, a positive consequence of this breeding strategy would be a substantial reduction in the mean value of *EXT*, which—together with *LIG*—represent undesired chemical components, in particular when wood is intended for paper or bioethanol production. Such a selection strategy would have a slightly detrimental effect on wood quality (it would result in a c. 3% loss in density and 7% in stiffness), but it would not compromise tree growth, as both *DBH* and *VOL* would remain nearly unaffected. Alternatively, due to the very high range in *EXT* observed among the studied trees (~29%), and given the genetic relationships among *CEL*, *LIG* and *EXT* (positive between *CEL* and *LIG* and negative between *EXT* and the other two), the best strategy for genetic improvement of the chemical composition of Scots pine juvenile wood suitable for the abovementioned purposes appears to be to target selection against *EXT* first, while focus on increasing *CEL* under the constraint of no increase in *LIG* later when both the mean and range of *EXT* have already been reduced to satisfactory levels.

4.5 | Practical implications and conclusion

FTIR spectroscopy, coupled with multivariate regression modeling, has proved to be a promising technique for Scots pine breeding as it can rapidly, inexpensively, non-destructively, and with good accuracy determine the chemical composition of wood in a large number of samples. Unlike wet lab protocols, it suffices with fractional amounts of wood material (~5 mg) and thus offers the possibility of skipping the extraction of increment cores from most trees included in the genetic evaluation. This would considerably reduce the stress load posed on trees, as increment core borers typically create big holes in their stems, which might jeopardize young trees' physical stability as well as increase the risk of fungal infections, in particular in trees that are to be maintained in their plantations or stands for a long time, e.g., until the rotation age. The resistograph appears to be an ideal tool in this regard, especially when wood density is already involved in the evaluation, as wood shavings produced during drill needle penetration can be collected and later utilized in the FTIR analysis.

The extent of the additive genetic variation observed among the studied progeny in the four major chemical compositional traits, i.e., the proportion of cellulose, hemicelluloses, lignin and extractives, in Scots pine juvenile wood along with their moderate genetic control indicates that the studied progeny tests possess good potential for future improvement via selective breeding. It seems most appropriate to commence the genetic improvement by decreasing and stabilizing the content of extractives among trees and then focus on increasing the ratio of cellulose to lignin, which is desirable for the paper producing industry as well as for the conversion of wood into biofuel products. Since extractives are moderately to highly heritable, it might also be of practical interest to investigate in more detail the predictability of individual extractive components to determine which one(s) stand behind their high correlation with FTIR absorbance intensities.

ACKNOWLEDGEMENTS

The authors thank Skogforsk (Forestry Research Institute of Sweden) for providing growth data from and access to the progeny tests; Sandu Seres, Ruth Ferguson and members of the Forest Genetics Lab at Umeå Plant Science Centre, SLU for help with field measurements and sample preparation; and András Gorzsás for advice on FTIR spectroscopy. Financial support from FORMAS, grant number 2014-00171, is greatly appreciated.

CONFLICT OF INTEREST

The authors declare no conflict of interest. The funder had no role in the design of the study, in the collection, analyses, or interpretation of data, in the writing of the manuscript or in the decision to publish the results.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available from Harvard Dataverse at https://doi.org/10.7910/DVN/KTHOXS.

ORCID

Tomáš Funda https://orcid.org/0000-0003-1275-9906

REFERENCES

Acquah, G. E., Via, B. K., Fasina, O. O., & Eckhardt, L. G. (2015). Non-destructive prediction of the properties of forest biomass for

chemical and bioenergy applications using near infrared spectroscopy. *Journal of Near Infrared Spectroscopy*, *23*(2), 93–102. https://doi.org/10.1255/jnirs.1153

Acquah, G. E., Via, B. K., Fasina, O. O., & Eckhardt, L. G. (2016). Rapid quantitative analysis of forest biomass using Fourier transform infrared spectroscopy and partial least squares regression. *Journal of Analytical Methods in Chemistry*, *2016*, 1–10. https://doi.org/10.1155/2016/1839598

Barnett, J., & Jeronimidis, G. (2003). *Wood quality and its biological basis*. Oxford, UK: Blackwell Publishing Ltd.

Binder, J. B., & Raines, R. T. (2009). Simple chemical transformation of lignocellulosic biomass into furans for fuels and chemicals. *Journal of the American Chemical Society*, *131*(5), 1979–1985. https://doi.org/10.1021/ja808537j

Bjarnestad, S., & Dahlman, O. (2002). Chemical compositions of hardwood and softwood pulps employing photoacoustic Fourier transform infrared spectroscopy in combination with partial least-squares analysis. *Analytical Chemistry*, *74*(22), 5851–5858. https://doi.org/10.1021/ac025926z

Brandel, G. (1990). *Volume functions for individual trees: Scots pine (Pinus sylvestris), Norway spruce (Picea abies) and birch (Betula pendula & Betula pubescens)*. Report 26, Swedish University of Agricultural Sciences. 180 pp. CABDirect database.

Bucur, V. (2006). *Acoustics of wood* (2nd ed.). Berlin, Heidelberg, Germany: Springer-Verlag.

Carter, P., Briggs, D., Ross, R. J., & Wang, X. (2005). *Acoustic testing to enhance western forest values and meet customer wood quality needs*. Paper presented at the Productivity of Western forests: A forest products focus. Gen. Tech. Rep. PNW-GTR-642., Pacific Northwest Research Station, Portland, OR, USA.

Chen, Z.-Q., Gil, M. R. G., Karlsson, B., Lundqvist, S.-O., Olsson, L., & Wu, H. X. (2014). Inheritance of growth and solid wood quality traits in a large Norway spruce population tested at two locations in southern Sweden. *Tree Genetics & Genomes*, *10*(5), 1291–1303. https://doi.org/10.1007/s11295-014-0761-x

Chen, Z.-Q., Karlsson, B., Lundqvist, S.-O., Gil, M. R. G., Olsson, L., & Wu, H. X. (2015). Estimating solid wood properties using Pilodyn and acoustic velocity on standing trees of Norway spruce. *Annals of Forest Science*, *72*(4), 499–508. https://doi.org/10.1007/s13595-015-0458-9

Cheverud, J. M. (1988). A comparison of genetic and phenotypic correlations. *Evolution*, *42*(5), 958–968. https://doi.org/10.2307/2408911

Conrad, A. O., & Bonello, P. (2016). Application of infrared and Raman spectroscopy for the identification of disease resistant trees. *Frontiers in Plant Science*, *6*, 1152. https://doi.org/10.3389/fpls.2015.01152

Costa e Silva, J., Wellendorf, H., & Pereira, H. (1998). Clonal variation in wood quality and growth in young sitka spruce (*Picea sitchensis* (Bong.) Carr.): Estimation of quantitative genetic parameters and index selection for improved pulpwood. *Silvae Genetica*, *47*(1), 20–33.

Cown, D. J., Young, G. D., & Burdon, R. D. (1992). Variation in wood characteristics of 20-year-old half-sib families of *Pinus radiata*. *New Zealand Journal of Forestry Science*, *22*(1), 63–76.

Cozzolino, D. (2014). Use of infrared spectroscopy for in-field measurement and phenotyping of plant properties: Instrumentation, data analysis, and examples. *Applied Spectroscopy Reviews*, *49*(7), 564–584. https://doi.org/10.1080/05704928.2013.878720

Eilers, P. H. C. (2004). Parametric time warping. *Analytical Chemistry*, *76*(2), 404–411. https://doi.org/10.1021/ac034800e

Ekeberg, D., Flaete, P. O., Eikenes, M., Fongen, M., & Naess-Andresen, C. F. (2006). Qualitative and quantitative determination of extractives in heartwood of Scots pine (*Pinus sylvestris* L.) by gas chromatography. *Journal of Chromatography A*, *1109*(2), 267–272. https://doi.org/10.1016/j.chroma.2006.01.027

El-Kassaby, Y. A., Mansfield, S., Isik, F., & Stoehr, M. (2011). In situ wood quality assessment in Douglas-fir. *Tree Genetics & Genomes*, *7*(3), 553–561. https://doi.org/10.1007/s11295-010-0355-1

Faix, O. (1992). Fourier transform infrared spectroscopy. In S. Y. Lin & C. W. Dence (Eds.), *Methods in lignin chemistry* (pp. 233–241). Berlin, Heidelberg, Germany: Springer. https://doi.org/10.1007/978-3-642-74065-7_16

Fengel, D., & Wegener, G. (1989). *Wood: Chemistry, ultrastructure, reactions*. Berlin, Germany: Walter de Gruyter.

Fries, A. (2012). Genetic parameters, genetic gain and correlated responses in growth, fibre dimensions and wood density in a Scots pine breeding population. *Annals of Forest Science*, *69*(7), 783–794. https://doi.org/10.1007/s13595-012-0202-7

Fries, A., Ericsson, T., & Gref, R. (2000). High heritability of wood extractives in *Pinus sylvestris* progeny tests. *Canadian Journal of Forest Research*, *30*(11), 1707–1713. https://doi.org/10.1139/cjfr-30-11-1707

Funda, T., Fundova, I., Gorzsás, A., Fries, A., & Wu, H. X. (2020). Predicting the chemical composition of juvenile and mature woods in Scots pine (*Pinus sylvestris* L.) using FTIR spectroscopy. *Wood Science and Technology*, *54*(2), 289–311. https://doi.org/10.1007/s00226-020-01159-4

Fundova, I., Funda, T., & Wu, H. X. (2018). Non-destructive wood density assessment of Scots pine (*Pinus sylvestris* L.) using Resistograph and Pilodyn. *PLoS One*, *13*(9), https://doi.org/10.1371/journal.pone.0204518

Fundova, I., Funda, T., & Wu, H. X. (2019). Non-destructive assessment of wood stiffness in Scots pine (*Pinus sylvestris* L.) and its use in forest tree improvement. *Forests*, *10*(6), 491. https://doi.org/10.3390/f10060491

Fundova, I., Hallingbäck, H. R., Jansson, G., & Wu, H. X. (2020). Genetic improvement of sawn-board stiffness and strength in Scots pine (*Pinus sylvestris* L.). *Sensors*, *20*(4), 1129. https://doi.org/10.3390/s20041129

Gapare, W. J., Wu, H. X., & Abarquez, A. (2006). Genetic control of the time of transition from juvenile to mature wood in *Pinus radiata* D. *Don. Annals of Forest Science*, *63*(8), 871–878.

Gebreselassie, M. N., Ader, K., Boizot, N., Millier, F., Charpentier, J.-P., Alves, A., … Segura, V. (2017). Near-infrared spectroscopy enables the genetic analysis of chemical properties in a large set of wood samples from *Populus nigra* (L.) natural populations. *Industrial Crops and Products*, *107*, 159–171. https://doi.org/10.1016/j.indcrop.2017.05.013

Gilmour, A. R., Gogel, B. J., Cullis, B. R., Welham, S. J., & Thompson, R. (2015). *ASReml user guide release 4.1*. Structural specification. Hemel Hempstead, UK: VSN International Ltd.

Gorzsás, A., & Sundberg, B. (2014). Chemical fingerprinting of *Arabidopsis* using Fourier transform infrared (FT-IR) spectroscopic approaches. In J. Sanchez-Serrano & J. Salinas (Eds.), *Arabidopsis protocols. Methods in molecular biology* (Vol. *1062*, pp. 317–352). Totowa, NJ: Humana Press.

Griffing, B. (1956). Concept of general and specific combining ability in relation to diallel crossing systems. *Australian Journal of Biological Sciences*, *9*, 463–493. https://doi.org/10.1071/BI9560463

Haapanen, M., Velling, P., & Annala, M. L. (1997). Progeny trial estimates of genetic parameters for growth and quality traits in Scots pine. *Silva Fennica*, *31*(1), 3–12. https://doi.org/10.14214/sf.a8506

Hayatgheibi, H., Forsberg, N. E. G., Lundqvist, S.-O., Mörling, T., Mellerowicz, E. J., Karlsson, B., … García-Gil, M. R. (2018). Genetic control of transition from juvenile to mature wood with respect to microfibril angle in Norway spruce (*Picea abies*) and lodgepole pine (*Pinus contorta*). *Canadian Journal of Forest Research*, *48*(11), 1358–1365.

Hayatgheibi, H., Fries, A., Kroon, J., & Wu, H. X. (2017). Genetic analysis of lodgepole pine (*Pinus contorta*) solid-wood quality traits. *Canadian Journal of Forest Research*, *47*(10), 1303–1313.

He, W., & Hu, H. (2013). Rapid prediction of different wood species extractives and lignin content using near infrared spectroscopy. *Journal of Wood Chemistry and Technology*, *33*(1), 52–64. https://doi.org/10.1080/02773813.2012.731463

Hein, P. R. G., & Chaix, G. (2014). NIR spectral heritability: A promising tool for wood breeders? *Journal of Near Infrared Spectroscopy*, *22*(2), 141–147. https://doi.org/10.1255/jnirs.1108

Hergert, H. L. (1971). Infrared spectra. In C. H. Ludwig & K. V. Sarkanen (Eds.), *Lignins: Occurrence, formation, structure and reactions*. New York: Wiley-Interscience.

Hodge, G. R., & Purnell, R. C. (1993). Genetic parameter estimates for wood density, transition age, and radial growth in slash pine. *Canadian Journal of Forest Research*, *23*(9), 1881–1891. https://doi.org/10.1139/x93-238

Hong, Z., Fries, A., & Wu, H. (2014). High negative genetic correlations between growth traits and wood properties suggest incorporating multiple traits selection including economic weights for the future Scots pine breeding programs. *Annals of Forest Science*, *71*(4), 463–472. https://doi.org/10.1007/s13595-014-0359-3

Isik, F., Li, B., & Frampton, J. (2003). Estimates of additive, dominance and epistatic genetic variances from a clonally replicated test of loblolly pine. *Forest Science*, *49*(1), 77–88.

Ivkovich, M., Namkoong, G., & Koshy, M. (2002). Genetic variation in wood properties of interior spruce. I. Growth, latewood percentage, and wood density. *Canadian Journal of Forest Research*, *32*(12), 2116–2127. https://doi.org/10.1139/x02-138

Jiang, W., Han, G., Via, B. K., Tu, M., Liu, W., & Fasina, O. (2014). Rapid assessment of coniferous biomass lignin-carbohydrates with near-infrared spectroscopy. *Wood Science and Technology*, *48*(1), 109–122. https://doi.org/10.1007/s00226-013-0590-3

Jozsa, L. A., & Middleton, G. R. (1994). *A discussion of wood quality attributes and their practical implications*. Vancouver, BC: Forintek Canada Corporation Western Laboratory, 42 pp.

Lepoittevin, C., Rousseau, J.-P., Guillemin, A., Gauvrit, C., Besson, F., Hubert, F., … Plomion, C. (2011). Genetic parameters of growth, straightness and wood chemistry traits in *Pinus pinaster*. *Annals of Forest Science*, *68*(4), 873–884. https://doi.org/10.1007/s13595-011-0084-0

Li, L., & Wu, H. X. (2005). Efficiency of early selection for rotation-aged growth and wood density traits in *Pinus radiata*. *Canadian Journal of Forest Research*, *35*(8), 2019–2029.

Loo, J. A., Tauer, C. G., & McNew, R. W. (1985). Genetic variation in the time of transition from juvenile to mature wood in loblolly pine (*Pinus taeda* L.). *Silvae Genetica*, *34*(1), 14–19.

Meder, R., Gallagher, S., Mackie, K. L., Bohler, H., & Meglen, R. R. (1999). Rapid determination of the chemical composition and density of *Pinus radiata* by PLS modelling of transmission and diffuse reflectance FTIR spectra. *Holzforschung*, *53*(3), 261–266. https://doi.org/10.1515/hf.1999.044

Nisula, L. (2018). *Wood extractives in conifers: A study of stemwood and knots of industrially important species*. Åbo, Finland: Åbo Akademi University Press. Retrieved from http://urn.fi/URN:ISBN:978-951-765-892-8

Nuopponen, M. H., Birch, G. M., Sykes, R. J., Lee, S. J., & Stewart, D. (2006). Estimation of wood density and chemical composition by means of diffuse reflectance mid-infrared Fourier transform (DRIFT-MIR) spectroscopy. *Journal of Agricultural and Food Chemistry*, *54*(1), 34–40. https://doi.org/10.1021/jf051066m

Pandey, A., Larroche, C., Ricke, S. C., Dussap, C.-G., & Gnansounou, E. (2011). *Biofuels: Alternative feedstocks and conversion processes*. Oxford, UK: Academic Press.

Paradis, N., Auty, D., Carter, P., & Achim, A. (2013). Using a standing-tree acoustic tool to identify forest stands for the production of mechanically-graded lumber. *Sensors*, *13*(3), 3394–3408. https://doi.org/10.3390/s130303394

Partanen, J., Harju, A. M., Venalainen, M., & Karkkainen, K. (2011). Highly heritable heartwood properties of Scots pine: Possibilities for selective seed harvest in seed orchards. *Canadian Journal of Forest Research*, *41*(10), 1993–2000. https://doi.org/10.1139/x11-116

Pearson, R. G., & Gilmore, R. C. (1980). Effect of fast growth rate on mechanical properties of loblolly pine. *Forest Products Journal*, *30*(5), 47–54.

Pettersen, R. C. (1984). The chemical composition of wood. In R. M. Rowell (Ed.), *The chemistry of solid wood* (Vol. *207*, pp. 57–126). Washington, DC: American Chemical Society.

Petty, J. A., MacMillan, D. C., & Steward, C. M. (1990). Variation of density and growth ring width in stems of Sitka and Norway spruce. *Forestry*, *63*(1), 39–49. https://doi.org/10.1093/forestry/63.1.39

Poletto, M., Zattera, A. J., & Santana, R. M. C. (2012). Structural differences between wood species: Evidence from chemical composition, FTIR spectroscopy, and thermogravimetric analysis. *Journal of Applied Polymer Science*, *126*, E336–E343. https://doi.org/10.1002/app.36991

Porth, I., Klapste, J., Skyba, O., Lai, B. S. K., Geraldes, A., Muchero, W., … Mansfield, S. D. (2013). *Populus trichocarpa* cell wall chemistry and ultrastructure trait variation, genetic control and genetic correlations. *New Phytologist*, *197*(3), 777–790. https://doi.org/10.1111/nph.12014

Pot, D., Chantre, G., Rozenberg, P., Rodrigues, J. C., Jones, G. L., Pereira, H., … Plomion, C. (2002). Genetic control of pulp and timber properties in maritime pine (*Pinus pinaster* Ait.). *Annals of Forest Science*, *59*(5–6), 563–575. https://doi.org/10.1051/forest:2002042

Räisänen, T., & Athanassiadis, D. (2013). *Basic chemical composition of the biomass components of pine, spruce and birch. Forest Refine. Finnish Forest Research Institute*.

Ramage, M. H., Burridge, H., Busse-Wicher, M., Fereday, G., Reynolds, T., Shah, D. U., … Scherman, O. (2017). The wood from the trees: The use of timber in construction. *Renewable & Sustainable Energy Reviews*, *68*, 333–359. https://doi.org/10.1016/j.rser.2016.09.107

SAS. (2008). *SAS/STAT® 9.2 user's guide*. Cary, NC: SAS Institute Inc.

Savitzky, A., & Golay, M. J. E. (1964). Smoothing + differentiation of data by simplified least squares procedures. *Analytical Chemistry*, *36*(8), 1627–1639. https://doi.org/10.1021/ac60214a047

Schimleck, L. R. (2008). Near infrared spectroscopy: A rapid, non-destructive method for measuring wood properties and its application to tree breeding. *New Zealand Journal of Forestry Science*, *38*(1), 14–35.

Schubert, C. (2006). Can biofuels finally take center stage? *Nature Biotechnology*, *24*(7), 777–784. https://doi.org/10.1038/nbt0706-777

Shupe, T. F., Choong, E. T., & Yang, C. H. (1996). The effects of silvicultural treatments on the chemical composition of plantation-grown loblolly pine wood. *Wood and Fiber Science*, *28*(3), 295–300.

Sjöström, E. (1993). *Wood chemistry: Fundamentals and applications*. San Diego, CA: Academic Press.

SPPBTC. (2003). *Protocol SCAN-CM 67:03: Wood chips for pulp production and pulp – Content of extractable lipophilic matter* (pp. 1–6). Stockholm, Sweden: Scandinavian Pulp, Paper and Board Testing Committee.

SPPBTC. (2009). *Protocol SCAN-CM 71:09. Pulps – Carbohydrate composition* (pp. 1–6). Stockholm, Sweden: Scandinavian Pulp, Paper and Board Testing Committee.

Stevanovic, T. (2016). Chemical composition and properties of wood. In N. Belgacem & A. Pizzi (Eds.), *Lignocellulosic fibers and wood handbook: Renewable materials for today's environment* (pp. 49–106). Hoboken, NJ: John Wiley & Sons.

Sticklen, M. B. (2010). Plant genetic engineering for biofuel production: Towards affordable cellulosic ethanol (Retraction of vol 9, pg 433, 2008). *Nature Reviews Genetics*, *11*(4), 308. https://doi.org/10.1038/nrg2777

Sykes, R., Isik, F., Li, B. L., Kadla, J., & Chang, H. M. (2003). Genetic variation of juvenile wood properties in a loblolly pine progeny test. *Tappi Journal*, *2*(12), 3–8.

Sykes, R., Li, B. L., Isik, F., Kadla, J., & Chang, H. M. (2006). Genetic variation and genotype by environment interactions of juvenile wood chemical properties in *Pinus taeda* L. *Annals of Forest Science*, *63*(8), 897–904. https://doi.org/10.1051/forest:2006073

TAPPI. (1991). *TAPPI UM 250: Acid-soluble lignin in wood and pulp. TAPPI test methods*. Atlanta, GA: Tappi Press.

TAPPI. (2002). *TAPPI T 222 om-02: Acid-insoluble lignin in wood and pulp. TAPPI test methods*. Atlanta, GA: Tappi Press.

Toivanen, T. J., & Alen, R. (2006). Variations in the chemical composition within pine (*Pinus sylvestris*) trunks determined by diffuse reflectance infrared spectroscopy and chemometrics. *Cellulose*, *13*(1), 53–61. https://doi.org/10.1007/s10570-005-9016-1

Ukrainetz, N. K., Kang, K.-Y., Aitken, S. N., Stoehr, M., & Mansfield, S. D. (2008). Heritability and phenotypic and genetic correlations of coastal Douglas-fir (*Pseudotsuga menziesii*) wood quality traits. *Canadian Journal of Forest Research*, *38*(6), 1536–1546. https://doi.org/10.1139/x07-234

Vandervoet, H. (1994). Comparing the predictive accuracy of models using a simple randomization test. *Chemometrics and Intelligent Laboratory Systems*, *25*(2), 313–323. https://doi.org/10.1016/0169-7439(94)85050-x

Wegner, T., Skog, K. E., Ince, P. J., & Michler, C. J. (2010). Uses and desirable properties of wood in the 21st century. *Journal of Forestry*, *108*(4), 165–173.

Wilhelmsson, L., & Andersson, B. (1993). *Breeding of Scots pine (Pinus sylvestris) and lodgepole pine (Pinus contorta ssp. latifolia)*. Paper presented at the Corrected reprint from S. J. Lee (Ed.). Progeny testing and breeding strategies, Proceedings of the Nordic group of tree breeding, Edinburgh.

Wu, H. X., Ivkovic, M., Gapare, W. J., Matheson, A. C., Baltunis, B. S., Powell, M. B., & McRae, T. A. (2008). Breeding for wood quality and profit in *Pinus radiata*: A review of genetic parameter estimates and implications for breeding and deployment. *New Zealand Journal of Forestry Science*, *38*(1), 56–87.

Xu, F., Yu, J. M., Tesso, T., Dowell, F., & Wang, D. H. (2013). Qualitative and quantitative analysis of lignocellulosic biomass using infrared techniques: A mini-review. *Applied Energy*, *104*, 801–809. https://doi.org/10.1016/j.apenergy.2012.12.019

Yang, K. C., & Benson, C. A. (1997). *Formation, distribution and its criteria for determining the juvenile-mature wood transition zone*. Paper presented at the CTIA/IUFRO International Wood Quality Workshop, Timber Management Toward Wood Quality and End-Product Value, Quebec City, Canada, 18–22 August 1997.

Zhang, S. Y., & Morgenstern, E. K. (1995). Genetic variation and inheritance of wood density in black spruce (*Picea mariana*) and its relationship with growth: Implications for tree breeding. *Wood Science and Technology*, *30*(1), 63–75. https://doi.org/10.1007/BF00195269

Zhou, C. F., Jiang, W., Cheng, Q. Z., & Via, B. K. (2015). Multivariate calibration and model integrity for wood chemistry using Fourier transform infrared spectroscopy. *Journal of Analytical Methods in Chemistry*, *2015*, 1–9. https://doi.org/10.1155/2015/429846

Zhou, C. F., Jiang, W., Via, B. K., Fasina, O., & Han, G. T. (2015). Prediction of mixed hardwood lignin and carbohydrate content using ATR-FTIR and FT-NIR. *Carbohydrate Polymers*, *121*, 336–341. https://doi.org/10.1016/j.carbpol.2014.11.062

Zhou, H., & Smith, I. (1991). Factors influencing bending properties of white spruce lumber. *Wood and Fiber Science*, *23*(4), 483–500.

Zhou, Z., Li, G., Huang, G., Chen, B., & Lin, Y. (2000). Genetic control of wood chemical composition and its implication for wood breeding of Masson pine. *Scientia Silvae Sinicae*, *36*(2), 110–115.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Funda T, Fundová I, Fries A, Wu HX. Genetic improvement of the chemical composition of Scots pine (*Pinus sylvestris* L.) juvenile wood for bioenergy production. *GCB Bioenergy*. 2020;12:848–863. https://doi.org/10.1111/gcbb.12723